Massively parallel *ab initio* simulations of semiconductor surfaces and nanotubes

J. Bernholc, E. L. Briggs, C. Bungaro, M. Buongiorno Nardelli,
J.-L. Fattebert, K. Rapcewicz, C. Roland, W. G. Schmidt and Q. Zhao
NC State University, Raleigh, NC 27695-8202

Abstract

We describe the results of extensive ab initio investigations of the properties of semiconductors and nanotubes. The following topics are discussed: (i) Ordering effects during surface incorporation of boron, which lead to the formation of ordered domains at half-monolayer coverage. These structures should lead to reduced impurity scattering in devices; (ii) Incorporation of Mg into GaN during growth, and in particular the conditions that would lead to maximum p-type doping; (iii) Optical fingerprints of surface structures for use in real-time feedback control of growth; and (iv) Calculations of the strength and transport properties of carbon nanotubes. All simulations were carried out using a novel, multigrid-based electronic structure method that enables very large-scale parallel simulations.

Introduction

The principal objective of our research is to predict and/or explain the properties of materials from first-principles, parameter-free calculations and to develop new methods for such calculations. The present work spans a substantial range: from studies of defects and impurities in bulk semiconductors and layered structures, to semiconductor growth, interface formation, and the theory of new, carbon-based materials. From a technical point of view, however, it consists mainly of ab initio molecular dynamics simulations. These simulations require substantial amounts of supercomputer time, but are very reliable and accurate. Apart from ab initio methods, we also use classical and tight-binding molecular dynamics codes for preliminary and follow-up research. These codes are sufficiently fast to run on our local workstations.

Much of our work focuses on a microscopic, atomic-level theory of complex semiconductor materials. The specific projects concern the electronic properties, mechanisms of growth, formation of interfaces, electron confinement in narrow quantum wells, and doping and defect formation mechanisms. Understanding of these phenomena on an atomic scale is of increasing importance due to the continuing shrinkage of semiconductor devices, which places increasingly stringent requirements on "perfection" at the atomic level. For very high speed devices, impurity scattering is a major problem, and one of our projects examined the ordering mechanisms in doping with B, which could substantially reduce electron scattering

phenomena. We have showed that several ordered structures are energetically degenerate, and that seemingly disordered B-doped surfaces can in fact consist of fairly large ordered domains, which should significantly reduce scattering.

For "new" semiconductor materials, e.g., those with wide band gaps, improvements in growth and doping methodologies are a major concern, since the quality of bulk material is not sufficient for advanced applications. Recently, there have been major breakthroughs in this area with the introduction of commercial high intensity blue LEDs from GaN and the production of the first long-lasting blue lasers. These developments will enable large gains in computer storage and full-color displays. Nitride semiconductors are also very suitable for high speed, high temperature devices. However, their p-type carrier density is only marginal for many of the proposed applications. We have thus investigated the doping mechanisms with Mg, the preferred p-type dopant, and identified conditions that could result in increased doping efficiency, well beyond the bulk solubility limit.

One of our main interests concerns generic mechanisms of growth of semiconductors. For advanced applications, controlled growth under industrial conditions is a major issue. Since such growth occurs in a gas ambient, the well-developed high-vacuum diagnostic techniques cannot be used. The only possible feedback sensors must be based on optical spectroscopy. However, the optical signal is often difficult to interpret and one needs to compute and identify the "fingerprints" of the various surface structures. Such calculations have recently become possible due to progress in theoretical methods and computer hardware. We present the first ab initio results for semiconductors in the InGaAsP family, used in lasers employed in optical communications. Better control over growth should lead to much denser, vertical cavity lasers, and improved high speed transistors.

Nanotubes are probably the most fascinating materials discovered in this decade. They are the strongest materials known and they also have unique electronic and emissive properties. We already determined the initial failure mechanisms under strain by approximate techniques, but in order to reliably predict the ultimate strength and failure rates an ab initio investigation of the barriers for atomic transformations is needed. This investigation is in progress, and the first results are described in the last section. Since mechanical transformation alter the conducting properties of nanotubes through plastic deformations, we are also investigating the quantum conductance of nanotubes by ab initio techniques, and present the first results using a new method for computing the conductance of large systems.

Inherent in our work is continuous improvement in methods and techniques, as well as development of new methods for large-scale calculations. All of the calculations described below were carried out using the real-space multigrid method that we developed [1]. It parallelizes very effectively and its speed increases linearly with the number of processors. Very recently, we developed a new, localized formulation of this method, which by using the so-called O(N) techniques should be able to substantially increase the size of the systems that can be studied [2]. It also leads to optimized localized orbitals, which we are using to compute the quantum conductance of nanotubes. Our formalism for calculating the quantum conductance [3] employs transfer matrix techniques and is also particularly suitable for studies of large systems.

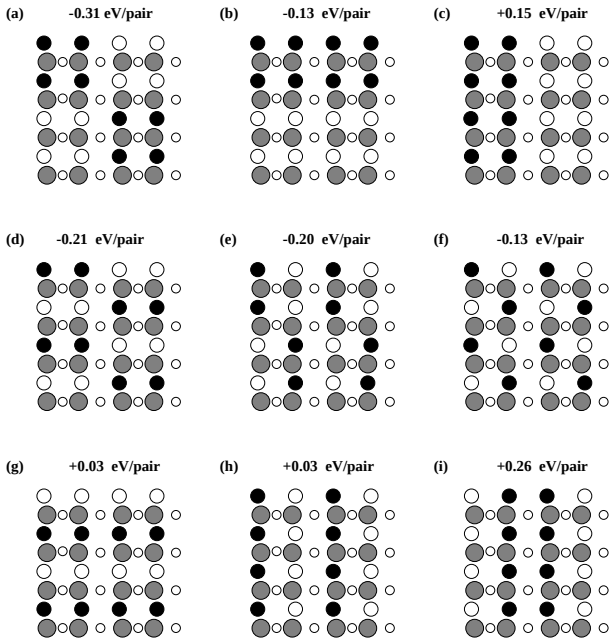

Figure 1: Schematic structures and energies for 0.5 ML surface doping.

Chemical and strain effects at Si(100):B

The segregation and ordering of boron impurities at high concentration levels near the Si(100) surface is due to a complex interplay between chemical and strain effects. Like the other acceptors in group III (Al, Ga, In), boron is favored to segregate to the surface, in order to saturate surface dangling bonds and to relieve the strain arising from its large atomic size mismatch with host Si atoms. However, unlike other group III acceptors, it prefers substitutional configurations in the first and second layers of the surface [4, 5]. With increasing doping, aggregation of boron impurities in the second layer of the surface becomes more and more favorable, up to a critical level corresponding to 0.5 ML of the impurities. Low-energy modes of ordering correspond to zigzag patterns of subsurface pairs. Above the critical doping level, the binding energies of ordered impurity configurations decrease sharply, and they become highly repulsive for a complete surface or subsurface impurity monolayer.

A number of experimental observations may be explained on the basis of our results. The equilibrium structure found at 0.5 ML doping has c(4x4) symmetry. However, both at low and high doping levels, several distinct modes of ordering are found to be essentially degenerate. This favors the formation of domain structures over long-range ordering of any one structure. Therefore, even at 0.5 ML coverage only (2x1) LEED patterns are observed, [6, 7, 8] while domains with c(4x4) and (4x4) periodicities are observed in STM [9]. If some of the metastable ordered structures can be stabilized, there would be a potential for engineering complex impurity patterns at the nanoscale.

One common feature of all the surfaces with subsurface boron atoms is the pronounced flattening and lowering of surface Si dimers with B neighbors, of up to 40-50% of the interlayer spacing. The lowered Si dimers are not visible in simulated STM images. However, the removal of such boron-bonded Si dimers to create surface divacancies is energetically very expensive, much more so than on the clean Si(100) surface [5]. Therefore, the strain fields

induced by heavy boron doping are not relieved by the creation of ordered patterns of surface Si divacancies. These results provide an alternative and more consistent explanation of the dark features extensively observed on B-doped Si(100) in STM investigations [9, 10, 11].

Mg incorporation at GaN surfaces

Controlled incorporation of impurities is critical to semiconductor science and technology. Specific dopant concentrations are necessary to achieve desired conductivities, recombination rates in light emitting devices or other electrical characteristics. In general, doping of bulk materials can be performed in three different ways: (i) by in-diffusion following growth, (ii) by ion implantation, and (iii) during growth. The in-diffusion of impurities is limited by the maximum solubility, which is a thermodynamic quantity that can be determined from either experiments or calculations. Ion implantation has no such limit, but in many materials the implantation damage is severe and cannot be annealed out, limiting the usefulness of this technique. During growth, the incorporation of an impurity occurs at the surface and depends sensitively upon the surface and its environment. For example, impurity energetics and consequently its solubility in the near-surface region can be significantly different from that in the bulk. This is because under typical growth conditions diffusion at the surface is much faster than in the bulk and the impurity density of the near-surface region can be frozen in as the film grows, leading to a concentration in the grown film different from that expected on the basis of its bulk solubility. Indeed, under these conditions it is the surface properties, instead of the bulk properties, that most strongly influence the impurity concentration. This is particularly true in a compound semiconductor, where many variables affect impurity incorporation during growth. Surface reconstruction patterns are often complex; in a given growth direction there are at least two possible surface terminations, and the surface structure varies as a function of the chemical potentials of the atomic constituents. The interplay of these effects leads to a complicated doping behavior, which, however, can be used to achieve a desired impurity concentration.

We have examined the incorporation of Mg in wurtzite GaN, which is an important paradigmatic case. There is a growing awareness that the two polarities of the (0001) GaN surface, nominally corresponding to the Ga and N faces, exhibit very different behavior. Indeed, striking differences have been observed in the morphology of the (0001) and $(000\bar{1})$ surfaces [12, 13, 14, 15, 16, 17, 18], and the strongly ionic nature of the GaN bond together with the low symmetry of the surface results in unusual reconstruction patterns that differ significantly from those observed in other III-V semiconductors. Since as-grown GaN exhibits unintentional n-type conductivity and controlled doping of GaN is necessary for optoelectronic devices, the achievement of good p-type conductivity has been a priority. However, Mg, which is the preferred p-type dopant, has a relatively large ionization energy. Consequently, high Mg concentrations are required to achieve the desired hole densities. In practice, the refined control of doping required to obtain reproducibly the particularly high Mg concentrations needed has been difficult to accomplish. Our calculations show that the interplay of surface orientation, reconstruction patterns and the availability of the species involved (as measured by their chemical potentials) determines the incorporation characteristics, and that intuition derived from bulk solubility considerations may be misleading [19]. For example, superior incorporation of Mg at the Ga-substitutional site (Mg_{Ga}) is expected

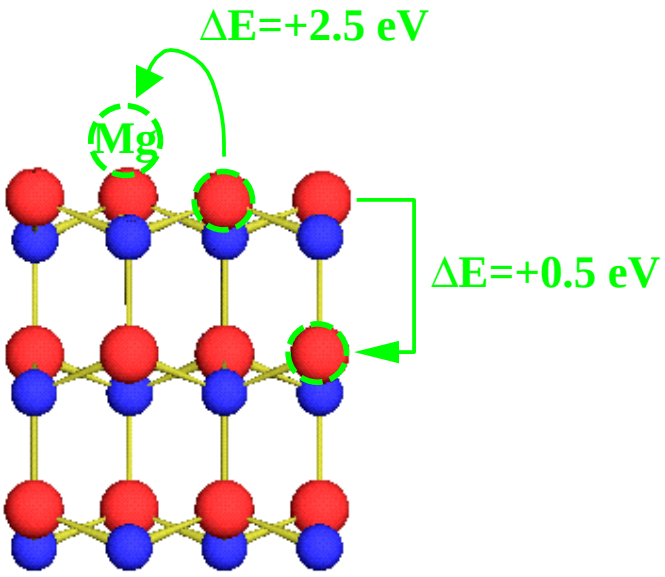

Figure 2: The energetics of Mg incorporation at the Ga-terminated surface.

under N-rich conditions, because of the decreased competition between Mg and Ga for the Ga site. However, this reasoning is a poor guide for determining the conditions for efficient incorporation. Indeed, we found that the growth conditions for optimal incorporation are *reversed* when the Mg is incorporated through a surface of different orientation (in this case, the film has different polarity). For the (0001) Ga-surface, we find that best incorporation occurs at N-rich and moderate Ga-rich conditions. On the contrary, for the $(000\overline{1})$ N-surface, in N-rich conditions Mg displays a strong tendency to segregate and superior incorporation occurs in a Ga-rich environment. When the surface is gallium terminated, best incorporation should occur. Growth procedures may therefore need to be tailored in a nontrivial fashion to the properties of the particular growth surface(s) in order to achieve effective doping with specific impurities.

Reflectance anisotropy spectra for large systems

The techniques of optical spectroscopy are evolving from fundamental semiconductor-surface studies to the control of semiconductor processing *in situ* and with real-time feedback. In particular, Reflectance Difference Spectroscopy (RDS), also known as Reflectance Anisotropy Spectroscopy (RAS), is a reliable *in situ* monitoring tool, applicable to ultrahigh vacuum as well as to gas-phase environments [22, 23, 24, 25]. It is particularly useful for the wide range of materials that are grown by metal-organic vapor phase epitaxy (MOVPE), where the well-developed electron- or ion-beam techniques cannot be used. However, first-principles calculations of RDS spectra are very expensive, because thick slabs and a large number of excited states must be used in computing the surface optical response. We have incorporated the formalism [20, 21] for calculating RDS into our parallel multigrid code. This has enabled us to perform, to our knowledge for the first time, a well-converged study of the optical response of realistic III-V(001) surface structures [26]. In particular, we were able to resolve the controversies regarding the InP(001)(2×4) surface structure [27, 28].

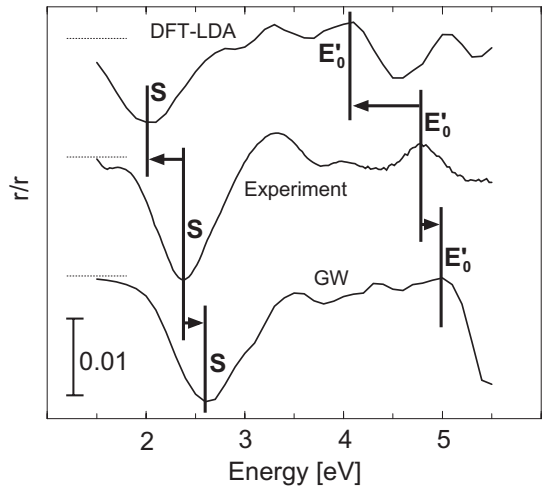

Figure 3: Calculated and measured RA spectra for the Ga-rich GaP surface. See text.

The comparison between theory and experiment, although hindered by the computational shortcomings of the DFT-LDA (see below), allowed for clear identification of the dominant surface features, or structural fingerprints, present under various growth conditions. This characterization should be very useful in real-time feedback and control of InP growth by MOVPE. However, the theoretical energy scale had to be "stretched" in order to match the experimental data. This is due to the well-known deficiencies of DFT-LDA in computing excited state properties.

In contrast to InP, the precise GaP(001) surface structure was essentially unknown prior to our study [29]. On the basis comprehensive total energy results we established the surface phase diagram. The cation-rich surface turned out to contain mixed dimers, in analogy to the corresponding InP surface. Our theoretical finding is corroborated by recent core-level results [30].

The RA spectra calculated for the energetically favored structures are strongly dependent on the surface geometry: The structures featuring cation-cation surface bonds show a pronounced negative anisotropy in the low energy region, with minima between 2.0 and 2.3 eV. The strength of that anisotropy is directly correlated to the number of Ga-Ga bonds along the [110] direction. On the other hand, we find that P-P dimers give rise to a relatively broad positive anisotropy between about 2.4 and 4.4 eV. These relations between the calculated RA and the surface stoichiometry agree very well with the experimental observations [29]. Furthermore, the results obtained for GaP are in line with the earlier findings for InP, supporting the existence of structural "fingerprints" in the optical spectra. This is an important outcome of our study with respect to the technological application of optical spectroscopy for semiconductor growth.

While we achieved a good qualitative description of the measured optical anisotropy, in particular with respect to the chemical trends, the calculated spectra still deviate significantly from experiment, in particular concerning the energetic positions of the peaks and

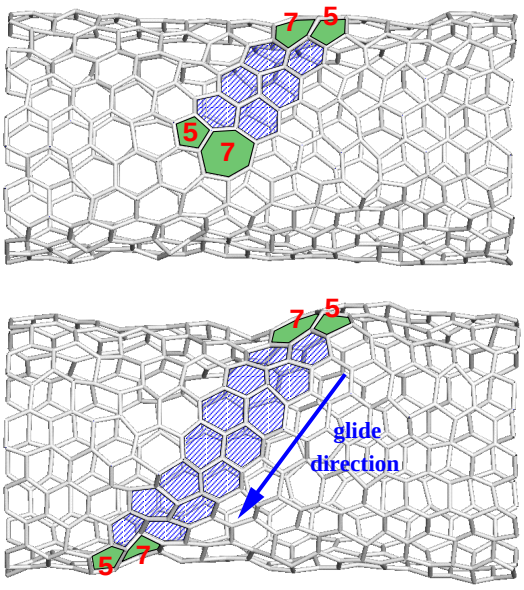

Figure 4: Plastic deformation in carbon nanotubes through the motion of 5-7 defects.

the line shape. In Fig. 3 we show the DFT-LDA calculation for the Ga-rich GaP surface in comparison with the measured curve. Highlighted in the figure is the pronounced negative anisotropy at low energy (denoted S). This feature is mainly related to optical transitions between filled Ga-Ga bonding states and empty dangling bond states located at surface Ga atoms. Its energetic position is underestimated in the DFT-LDA calculation by about 0.3 eV with respect to the measured spectrum [29]. A much stronger underestimation of about 0.7 eV occurs for the essentially bulk-related features at higher photon energies. This non-uniform shift of the calculated spectra cannot be remedied by a simple scissor-operator approach. We applied an approximate GW-approach, based on a model dielectric function in conjunction with an approximate treatment of local-field effects and the dynamical screening [31, 32]. The resulting RA spectrum is also presented in Fig. 3. Compared to the DFT-LDA results we observe considerable improvements: The most important effect of the quasiparticle corrections is a non-uniform shift of the spectrum towards higher photon energies. We also observe a stronger shift for the bulk-related features than of the surface signatures.

Nanotubes

Nanotubes are extremely unusual in their ability to withstand extreme deformations, and are thus particularly suitable to mechanical applications. We have previously shown that nanotubes are amazingly flexible and this prediction has already been confirmed experimentally [33]. A further prediction, based on classical MD simulations, is that nanotubes deform plastically under strain by a series of bond rotations [34]. This mechanism involves a formation of a 5-7-7-5 defect and its dissociation into two 5-7 pairs. Further strain results in the motion of 5-7 defects and thus plasticity, see Fig. 4. However, for a precise prediction of the value of the maximum tensile stress at various temperatures it is necessary to have accurate

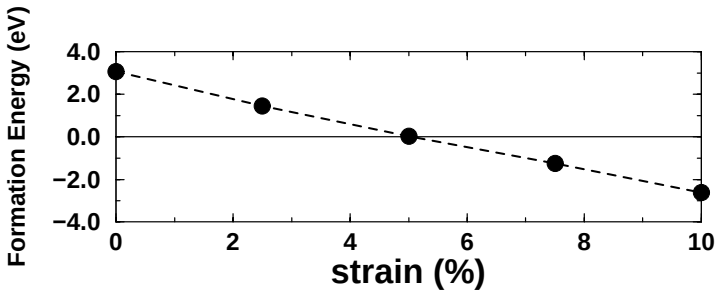

Figure 5: Formation energy of the (5-7-7-5) defect for the (5,5) tube at different strains.

values of both the formation and activation energies of these processes as a function of strain. We have already computed the formation energies of the 5-7-7-5 defect for a (5,5) armchair tube as function of strain, see Fig. 5. Additional calculations for nanotubes with different helicities and diameters are in progress, as well as calculations of the activation energies.

Furthermore, for a material at thermal equilibrium, the number of defects of a particular type given by $N_{\rm sites}e^{(-G_F/kT)}$, where $N_{\rm sites}$ is the number of potential sites and G_F is the Gibbs free energy of formation of the defect. The entropic contributions are usually a small part of G_F; one can thus use the energy of formation to obtain a lower bound. For single-walled tubes, which are grown at ~ 1500 K, the above formula suggests that even at zero strain there might be a point defect in nanotubes every few tenths of a mm. This would certainly limit the ultimate strength of nanotubes and thus their eventual uses.

Nanotubes can be either metallic or semiconducting, depending on their helicity. Metal-semiconductor junctions can easily be formed, raising the prospect of nanoscale all-carbon electronic devices [35, 34]. The problem of calculating quantum conductance in carbon nanotubes has already been addressed with a variety of techniques that reflect the various approaches in the theory of quantum transport in ballistic systems. Most of the existing calculations use a simple π-orbital tight-binding Hamiltonian that describes the electronic structure of a carbon nanotube via a single nearest-neighbor hopping parameter. This model provides a quite good qualitative description of the bands near the Fermi level and, given its simplicity, it has become the model of choice in a number of theoretical investigations. However, although qualitatively useful to interpret experimental results, this simple Hamiltonian lacks the accuracy that ab initio methods are able to provide.

Recently, we have developed a general scheme for calculating quantum conductance that is particularly suitable for realistic calculations of electronic transport properties in extended systems [3]. It is very flexible and applicable to any system described by a Hamiltonian with a localized orbital basis; the only quantities that enter into the present formulation are the matrix elements of the Hamiltonian operator, with no need for the explicit knowledge of the electron wave functions for the multichannel expansion. The last fact makes the numerical calculations particularly efficient also for systems described by multi-orbital localized-basis Hamiltonians.

In order to obtain the highly-localized ab initio orbitals necessary for the efficient application of this scheme, we have used a recently developed O(N)-like method for electronic structure calculations [2]. This method uses a non-orthogonal grid-based orbital description of the Kohn-Sham equations and multigrid acceleration to obtain relatively few variationally

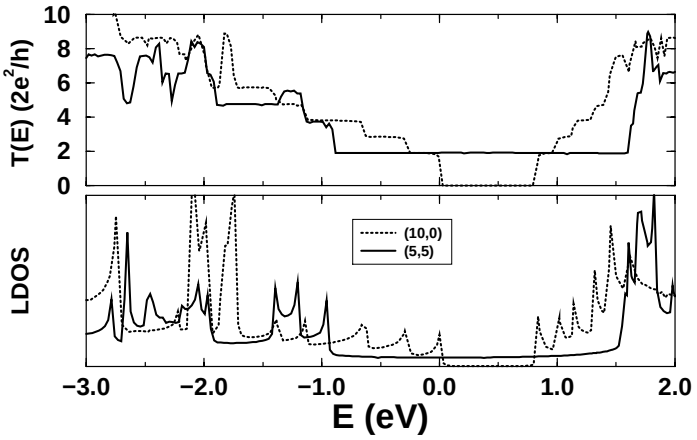

Figure 6: Quantum conductance and density of states for a metallic (5,5) and a semiconducting (10,0) carbon nanotube from ab initio calculations. The calculations used four optimized orbitals per atom.

optimized orbitals that minimize the total energy without compromising the accuracy. Using these techniques we have already computed the quantum conductance for two test systems, one metallic and one semiconducting. (Since nanotubes can be doped, understanding of the low energy region of semiconducting nanotubes is also important.) Our first results are shown in Fig. 6. Calculations for tubes with strain-created defects are in progress, with the aim of identifying the electrical signatures of such defects.

References

[1] E. L. Briggs, D. J. Sullivan and J. Bernholc, Phys. Rev. B **52**, R5471 (1995); Phys. Rev B **54**, 14362 (1996).

[2] J.-L. Fattebert *et al.*, to be published.

[3] M. Buongiorno Nardelli, to be published.

[4] J. E. Northrup et al., Phys. Rev. B **44**, 13799 (1991).

[5] M. Ramamoorthy, E. L. Briggs and J. Bernholc, Phys. Rev. Lett. **81**, 1642 (1998).

[6] R. L. Headrick, B. E. Weir, A. F. J. Levi, D. J. Eaglesham and L. C. Feldman, Phys. Rev. Lett. **57**, 2779 (1990).

[7] B. E. Weir, R. L. Headrick, Q. Shen, L. C. Feldman, M. S. Hybertsen, M. Needels, M. Schluter and T. R. Hart, Phys. Rev. B **46**, 12861 (1992).

[8] R. Cao, X. Yang and P. Pianetta, J. Vac. Sci. Tech. B **11**, 1455 (1993).

[9] Y. Wang, R. J. Hamers and E. Kaxiras, Phys. Rev. Lett. **74**, 403 (1995).

[10] Z. Zhang et al., J. Vac. Sci. Tech. B **14**, 2684 (1996).

[11] T. Komeda and Y. Nishioka, Appl. Phys. Lett. **71**, 2277 (1997).

[12] F. A. Ponce, D. P. Bour, W. T. Young, M. Saunders and J. W. Steeds, Appl. Phys. Lett. **69**, 337 (1996).

[13] B. Daudin, J. L. Roviére and M. Arlery, Appl. Phys. Lett. **69**, 2480 (1996).

[14] Z. Liliental-Weber, Y. Chen, S. Ruvimov and J. Washburn, Phys. Rev. Lett. **79**, 2835 (1997).

[15] M. M. Sung, J. Ahn, V. Bykov, J. W. Rabalais, D. D. Koleske and A. E. Wickenden, Phys. Rev. B **54**, 14652 (1996).

[16] E. S. Hellman, D. N. E. Buchanan, D. Wiesmann and I. Brener, MRS Internet J. Nitride Semicond. Res. **1**, 16 (1996).

[17] K. Rapcewicz, M. Buongiorno Nardelli and J. Bernholc, Phys. Rev. B **56**, R12725 (1997).

[18] A. R. Smith, R. M. Feenstra, D. W. Greve, J. Neugebauer and J. E. Nothrup, Phys. Rev. Lett. **79**, 3934 (1997); A. R. Smith, R. M. Feenstra, D. W. Greve, M.-S. Shin, M. Skowronski, J. Neugebauer and J. E. Nothrup, Appl. Phys. Lett. **72**, 2114 (1998).

[19] C. Bungaro, K. Rapcewicz and J. Bernholc, Phys. Rev. B **59**, 9771 (1999).

[20] R. Del Sole, Solid State Commun. **37**, 537 (1981).

[21] F. Manghi, R. Del Sole, A. Selloni, and E. Molinari, Phys. Rev. B **41**, 9935 (1990).

[22] D.E. Aspnes and A.A. Studna, Phys. Rev. Lett. **54**, 1956 (1985).

[23] D.E. Aspnes, Surf. Sci. **307-309**, 1017 (1994).

[24] W. Richter and J.T. Zettler, Appl. Surf. Sci **101**, 465 (1996).

[25] *Epioptics. Linear and Nonlinear Optical Spectroscopy of Surfaces and Interfaces* Ed. by J.F. McGilp, D. Weaire, and C.H. Patterson, (Springer-Verlag, Berlin, 1995).

[26] W.G. Schmidt, E.L. Briggs, J. Bernholc, and F. Bechstedt, Phys. Rev. B **59**, 2234 (1999).

[27] D. Pahlke, J. Kinsky, C. Schultz, M. Pristovsek, M Zorn, N Esser, W. Richter, Phys. Rev. B **56**, R1661 (1997); K.B. Ozanyan, P.J. Parbrook, M. Hopkinson, C.R. Whitehouse, Z. Sobiesierski, D.I. Westwood, J. Appl. Phys. **82**, 474 (1997); M. Zorn, T.Trepk, J.T. Zettler, B. Junno, C. Meyne, K. Knorr, T. Wethkamp, M. Klein, M. Miller, W. Richter, L. Samuelson, Appl. Phys. A **65**, 333 (1997); J. Kinsky, C. Schultz, D. Pahlke, A.M. Frisch, T. Herrmann, N. Esser, W. Richter, Appl. Surf. Sci. **123**, 228 (1998).

[28] C.D. MacPherson, R.A. Wolkow, C.E.J. Mitchell, A.B. McLean, Phys. Rev. Lett. **77**, 691 (1996); M. Shimomura, N. Sanada, Y. Fukuda, P.J. Moller, Surf. Sci. **359**, L451 (1996).

[29] A.M. Frisch *et al.* submitted to Phys. Rev. B.

[30] N.Sanada *et al.* Surf. Sci. **419**, 120 (1999).

[31] F. Bechstedt *et al.*, Sol. State Commun. *84*, 765 (1992).

[32] W.G. Schmidt *et al.*, in preparation.

[33] S. Iijima, C. Brabec, A. Maiti and J. Bernholc, J. Chem. Phys. **104**, 2089 (1996); M.R. Falvo, G.J. Clary, R.M. Taylor II, V. Chi, F.P. Brooks Jr, S. Washburn and R. Superfine, Nature, **389**, 582 (1997).

[34] M. Buongiorno Nardelli, B. I. Yakobson and J. Bernholc, Phys. Rev. B **57**, R4277 (1998).

[35] B.I. Dunlap, Phys. Rev. B **46**, 1933 (1992); L. Chico, V.H. Crespi, L.X. Benedict, S.G. Louie and M.L. Cohen, Phys. Rev. Lett. **76**, 971 (1996).